CUSTOMER SEGMENTATION

The Customer Segmentation is performed using K-Means clustering on a dataset combining customer and transaction details. The features used were region, total quantity purchased, transaction value, and transaction count, all standardized for consistency. A 2D scatter plot was used to visualize the 7 clusters. Clear distinctions between some clusters were observed, while others had slight overlaps.

❖ **Number of Clusters Formed**

The segmentation process resulted in **7 clusters**, which effectively grouped customers based on their spending patterns, transaction frequency, and region. Each cluster represents a distinct customer segment with specific characteristics and behaviors.

❖ **Davies-Bouldin Index (DB Index)**

The DB Index for the 7-cluster configuration was **0.99**, indicating strong performance in terms of cluster compactness and separation. This metric confirms that the clusters are well-defined with minimal overlap.

❖ **Other Relevant Clustering Metrics**

Silhouette Score : The average silhouette score **0.3068** for the configuration showed good intra-cluster similarity and inter-cluster separation.

❖ **Insights**

- ➢ **Cluster 1:** High-frequency, high-value customers - excellent for loyalty and retention efforts.
- ➢ **Cluster 2:** Low-transaction customers with low spend - potential for re-engagement campaigns.
- ➢ **Cluster 3:** Region-specific moderate spenders - opportunity for targeted regional strategies.
- ➢ **Cluster 4:** Consistent, moderate spenders - steady contributors to revenue.
- ➢ **Clusters 5, 6, 7:** Niche groups defined by specific combinations of spending and frequency - ideal for personalized marketing strategies.